

# **Improved Assessment of Personality Disorders that are Security Risks<sup>1</sup>**

*Eric L. Lang, Ph.D.*

*Defense Personnel Security Research Center (PERSEREC), Monterey, CA 93940, USA*

*Eric.Lang@osd.pentagon.mil*

*Olga G. Shechter, Ph.D.*

*Northrop Grumman Technical Services (at PERSEREC)*

*Olga.Shechter.ctr@osd.pentagon.mil*

## **PROBLEM AND OVERVIEW**

A key security challenge facing government agencies is prevention of insider threat attacks associated with trusted government employees, military personnel, and contractors. These attacks can take on a variety of forms, including espionage, terrorism and violence, compromise of information technology networks, and security incidents involving failure to adequately protect sensitive information and systems. Certain clinical personality disorders can increase the likelihood of insider threat; however, these disorders are difficult to diagnose with routine assessment tools that rely on the subject's self-report. A further complicating factor is that there is insufficient information on which disorders could lead to compromised judgment, reliability, and trustworthiness. In the wake of tragic insider events, such as the 2009 Fort Hood Massacre, 2010 "WikiLeaks" compromise of sensitive information, and 2001 espionage conviction of Robert Hanssen, improving the identification of insiders with personality disorders associated with security and safety risks is critical.

In order to address these challenges, the Defense Personnel Security Research Center (PERSEREC) initiated a three-phase research project aimed at ensuring that individuals with risky personality disorders are identified correctly and handled appropriately by clinicians, selection, and personnel security staff. Phase I examined which personality disorders pose the greatest risk to personnel security by engaging security risk specialists, senior adjudicators from the intelligence community, in an applied research initiative designed to create a security risk prototype (Godes & Lang, 2009). A Phase II field test extended this research by evaluating whether an improved personality disorder screening tool (the Shedler-Westen Assessment Procedure [SWAP]) and a metric of security risk developed from Phase I (Dispositional Indicators of Risk Exposure [DIRE]), have utility for clinicians who routinely evaluate personnel in a high-risk program that involves access to dangerous materials (Shechter & Lang, in prep). Finally, Phase III examined the criterion-related validity of DIRE by using it to predict broader safety and suitability risks, such as violent behavior and employment problems.

---

<sup>1</sup> The research described in this paper was presented by Eric Lang at the International Applied Military Psychology Symposium (IAMPS) in Vienna, Austria on May 25, 2011. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of PERSEREC or the USA Department of Defense.

## **PERSONALITY DISORDER ASSESSMENT CHALLENGES**

There are several challenges to correctly diagnosing individuals with personality disorders. First, afflicted individuals are frequently unaware of their symptoms, as certain personality disorders (i.e., narcissistic personality disorder) do not disrupt intellectual and perceptual functioning. Consequently, individuals' self-reported information about their symptoms is often incomplete or inaccurate. Second, individuals who are aware of their symptoms may be reluctant to report them, (e.g., in a job selection situation, because they fear rejection). Third, commonly used personality disorder assessment tools (e.g., Minnesota Multiphasic Personality Inventory Restructured Form [MMPI-2-RF]), Tellegen & Ben-Porath, 2008; Personality Assessment Inventory [PAI], Morey, 1991) rely on the subject's self-report, which is likely to be inaccurate in situations where the individual is not aware of his or her symptoms or is engaged in deception. In addition, although a well-trained clinician may be able to accurately diagnose severe personality disorders based on a clinical interview, such diagnoses often yield subjective descriptions that can vary across clinicians and lack standardized quantitative results.

## **SWAP**

The SWAP overcomes the described assessment challenges by not relying on the subject's self-report. It is based on a premise that clinicians can make reliable observations and inferences about personality organization if they are given a suitable technology for harnessing and quantifying their judgments (Shedler & Westen, 1998). Procedurally, clinicians first complete a narrative-based Clinical Diagnostic Interview (CDI). The CDI takes approximately 2.5 hours and collects information about the subject's personality organization, including his or her childhood, education and work history, and relationship history with family members and significant others (Bradley, Hilsenroth, Guarnaccia, & Westen, 2007). After administering the CDI, the clinician rates 200 personality-descriptive statements about the subject (e.g., "tends to get into power struggles") on a scale from 0 (inapplicable to the individual) to 7 (highly descriptive of the individual). The SWAP ratings, which take approximately 45 minutes, can also be completed without a CDI, if the clinician has comparable information on a subject (e.g., based on extensive prior clinical contact with the individual).

The quantitative interpretation of the SWAP data is performed automatically by the SWAP-200 program that contains built-in algorithms for computing the subject's scores on all Axis II personality disorders from *DSM-IV* (Paranoid, Schizoid, Schizotypal, Antisocial, Borderline, Histrionic, Narcissistic, Avoidant, Dependent, Obsessive, Depressive, Passive-Aggressive) as well as scores on clinical personality syndromes that surfaced in past research and clinical practice that are more typical of the range of pathology observed in actual patients (e.g., Psychopathy). In addition, the SWAP computes two indices of overall mental health (Healthy Functioning and Personality Health Index). Extensive research evidence indicates that SWAP is a reliable and valid instrument for assessing personality disorders (e.g., Shedler & Westen, 2004) and it has been translated into 22 languages. As an example of relevance to personnel safety and security applications, prospective research indicated that SWAP scales are a better predictor of violent behavior than other commonly used measures of personality disorders (Marin de Avellan, 2010).

## **PHASE I: IDENTIFYING PERSONALITY DISORDERS THAT ARE SECURITY RISKS**

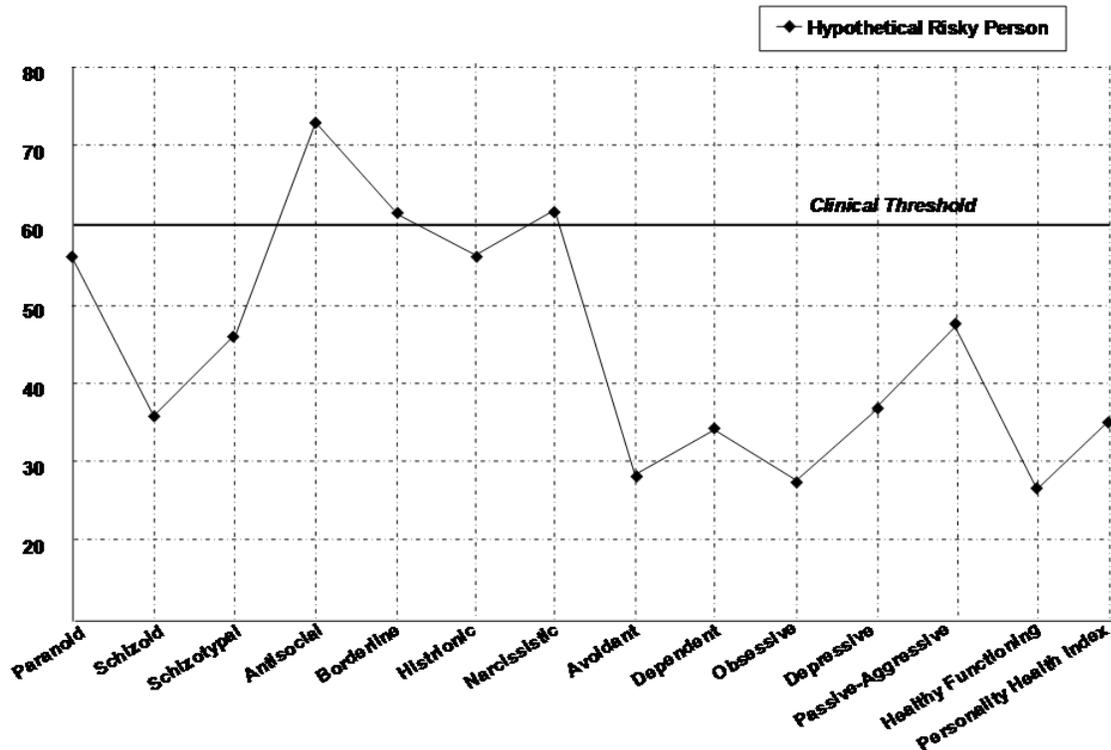
### **Goal and Procedure**

The goal of Phase I was to use the SWAP methodology to address the dearth of research on which personality disorders are associated with the highest level of security risk. Twenty experienced personnel security adjudicators from four separate government adjudication centers were asked to rate each SWAP item based on how much it applied to a hypothetical risky individual capable of endangering the safety of others, compromising important systems, or otherwise undermining national security. The 4,000 ratings (200 items x 20 adjudicators) were subsequently used to generate a personality-behavioral prototype of security risk and to develop the DIRE scale.

### **Results**

Reliability analyses showed that there was a high degree of agreement across adjudicators' ratings regarding which characteristics represent the greatest security risks (Cronbach's alpha = .92). Because inter-rater reliability was high, adjudicators' ratings were next subjected to a Q-factor analysis to identify the number of personnel security risk groupings or factors that best describe the structure of the data. A close examination of factor loadings and factor scores by Dr. Jonathan Shedler, a co-author of the SWAP, showed psychopathy, malignant narcissism, and borderline personality organization to be key personnel security risk factors. An aggregate SWAP profile of Axis II *DSM-IV* personality disorder scores shown in Figure 1, created by averaging the ratings of 20 adjudicators, confirmed this finding.

The diagnostic labels of the three disorders that exceed the clinical threshold of 60 in Figure 1 are slightly different from the labels that were assigned to them based on the Q-analysis. The *DSM-IV*'s classification of personality disorders does not adequately capture the richness of personality syndromes that are seen in clinical practice and that surfaced in adjudicators' ratings. For example, the construct of malignant narcissism differs from the *DSM-IV* construct of narcissism because it also incorporates psychopathic traits. Psychopathy is a broader construct than antisocial personality disorder. Unlike antisocial personality disorder (a *DSM-IV* construct), which is defined by persistent behavioral violations of social norms, psychopathy (not listed in *DSM-IV*) also encompasses interpersonal and affective symptoms such as a lack of remorse for harm or injury caused to others. Finally, borderline personality organization is a broader construct than borderline personality disorder, and individuals with this personality type frequently experience volatile emotions and are prone to self-harm. See Godes & Lang (2009) for a more thorough description of the three high-risk personality disorders.



**Figure 1 Axis II Personality Disorder SWAP Profile Produced by Averaging Adjudicators' Ratings**

### **Dispositional Indicators of Risk Exposure (DIRE) Subscale**

The Dispositional Indicators of Risk Exposure (DIRE) scale is a metric of security risk that was developed on the basis of adjudicators' ratings of SWAP items most descriptive of a hypothetical risky person. DIRE, a subscale of the SWAP, assesses a constellation of personality characteristics associated with psychopathy, malignant narcissism, and borderline personality organization. When these disorders occur in combination, the level of security risk increases substantially (Shedler, 2009). DIRE quantifies this degree of risk by measuring the match between a subject's personality and the high-risk personality prototype identified in Phase I research. DIRE was developed primarily for use with persons applying for, or occupying, sensitive positions requiring a security clearance or access to sensitive materials (e.g., nuclear weapons, biological toxins). The national security version of the SWAP-200 software, available from the SWAP commercial publisher, computes a DIRE score in addition to typical SWAP personality scores.

## **PHASE II: SWAP/DIRE FIELD TEST**

### **Goal and Procedure**

After determining which personality disorders were the greatest security risks in Phase I, the objective of the Phase II field test was to investigate whether the SWAP and DIRE have added

utility for clinicians who routinely evaluate personnel for a high-risk Department of Energy's Human Reliability Program (DoE HRP). HRP is a safety and security program designed to ensure that personnel who occupy positions affording access to dangerous materials, programs, and devices meet the highest standards of reliability and mental and physical health.

In October 2009, five DoE clinicians were trained by Dr. Shedler in all aspects of the SWAP method, administration of the CDI, and use of the DIRE scale. Clinicians also had a chance to practice using the SWAP software and procedure by evaluating a highly familiar case that was still fresh in their memory. Following the training, clinicians used SWAP/DIRE for a period of 4 months to evaluate the mental health fitness of 10 HRP candidates and 16 problematic employees. The candidates were personnel who held a DoE Q clearance (equivalent to Department of Defense's [DoD] Top Secret) and were now applying for HRP certification. Problematic employees were either existing HRP personnel who were being closely monitored because of prior concerns (employees of concern), or existing HRP personnel involved in a behavioral problem or incident of concern on the job and were being referred to the clinical team for further evaluation (new referrals for cause). Within 24 hours of each SWAP administration, clinicians filled out an effectiveness rating sheet assessing the perceived utility of SWAP/DIRE. In March 2010, clinicians participated in 1-1.5 hour phone debriefing interviews with PERSEREC research staff regarding their experience with SWAP/DIRE.

## **Results**

Analysis of the interview data indicated that clinicians found all components of the SWAP method useful for evaluating new HRP candidates and current HRP personnel they were either monitoring more closely or seeing as new referrals for cause. SWAP's CDI allowed the clinicians to establish a positive rapport with the subjects up front, which they felt was essential for instilling feelings of trust in the personnel and encouraging them to turn to the clinical team in the future if they begin experiencing problems on the job and/or in their personal life. This was an unexpected finding as several clinicians expressed concerns prior to the start of the field test that subjects may find the CDI questions overly personal. The CDI added approximately 45 minutes to the clinicians' standard interview.

Clinicians reported that the SWAP was the best instrument they were aware of for diagnosing personality disorders because, unlike the other tools, it enabled them to standardize and quantify their judgments and observations instead of relying on the subject's self-report of their personality characteristics. Clinicians also reported that they became faster at completing the SWAP ratings with each subsequent use of the instrument, averaging between 30-45 minutes per administration.

Because individuals with high DIRE scale scores are uncommon, clinicians had limited experience with this scale during the field test. Clinicians liked DIRE's relevance to security risk and its utility for reinforcing their recommendations regarding HRP eligibility, because DIRE scores represent a consensus of senior adjudicators regarding security risks. In contrast with their existing assessment tools, the SWAP metrics of psychological health allowed clinicians to better juxtapose an individual's personality strengths with his or her weaknesses when making a determination regarding the high-risk program fitness.

## **PHASE III: EXPLORING DIRE CRITERION-RELATED VALIDITY**

### **Goal and Procedure**

The objective of Phase III was to evaluate potential extensions of the SWAP/DIRE associations beyond traditional personnel security outcomes and into the domains of suitability and safety, e.g., by examining whether DIRE can predict violent behavior and troubled employment history. Prevention of violent behavior is crucial to national security, public safety, and anti-terrorism, especially in the wake of such events as the 2009 Fort Hood Massacre and the 2010 shooting rampage in Afghanistan carried out by a disgruntled contracted linguist (Department of Defense Independent Review, 2010; Lardner, 2011).

Data for Phase III came from a Shedler and Westen initiative funded by the National Institutes of Mental Health. A random national sample of 1,201 clinicians, drawn from membership registers of the American Psychiatric Association and American Psychological Association, was asked to randomly select, and then assess using the SWAP, one patient in their care with personality-related issues causing distress. After completing their SWAP ratings, clinicians filled out a checklist of the criteria for all *DSM-IV* Axis II disorders, and a Clinical Data Form assessing key behavioral outcomes about the patient, such as history of arrests and violent behavior, employment troubles, etc. The resultant sample consisted of 1,201 patients (53.2% female; 82.7% Caucasian; mean age 42.3 years; range of all socioeconomic classes). During the data analytic phase for PERSEREC, a DIRE score was computed for each patient, after which it was correlated with the behavioral risk composite variables constructed from the individual items on the Clinical Data Form.

### **Results**

The major findings are summarized in Table 1. DIRE was a significant predictor of adult maladaptive functioning, employment trouble, mental stability, forensic risk, and childhood adolescent misbehavior. Moreover, as shown in Table 2, DIRE also outperformed the individual measures of the three *DSM-IV* Axis II disorders that form its backbone (antisocial personality disorder, malignant narcissism, and borderline personality organization) in predicting adult maladaptive functioning and employment trouble. Borderline personality disorder was the best predictor of mental stability, and antisocial personality disorder was the best predictor of forensic risk and childhood and adolescent misbehavior. This finding is not surprising considering the fact that these three outcomes bear less relevance to suitability and safety risks than do adult maladaptive functioning and employment trouble. Taken together, these Phase III results establish DIRE as a valid predictor of suitability and safety risks, in addition to personnel security risks, and a superior predictor when compared to individual personality disorder scales.

**Table 1**  
**Correlations between DIRE and Behavioral Risk Composite Variables**

| <i>Variable Name</i>             | <i>Correlation with DIRE<sup>1</sup></i> | <i>% of Variance Explained by DIRE</i> |
|----------------------------------|--|--|
| Adult Maladaptive Functioning    | .64                                      | 41%                                    |
| Employment Trouble               | .49                                      | 24%                                    |
| Mental Stability                 | .34                                      | 12%                                    |
| Forensic Risk                    | .46                                      | 21%                                    |
| Childhood/Adolescent Misbehavior | .53                                      | 28%                                    |

<sup>1</sup> All correlations were statistically significant at  $p < .001$ .

**Table 2**  
**Comparing DIRE with Individual Personality Disorder Measures in Predicting Behavioral Risk Composite Variables**

| <i>Variable Name</i>             | <i>Correlation with DIRE<sup>1</sup></i> | <i>Antisocial Personality Disorder</i> | <i>Narcissistic Personality Disorder</i> | <i>Borderline Personality Organization</i> |
|----------------------------------|--|--|--|--|
| Adult Maladaptive Functioning    | .64                                      | .52                                    | .18                                      | .51  |
| Employment Trouble               | .49                                      | .42                                    | .18                                      | .35  |
| Mental Stability                 | .34                                      | .21                                    | .02                                      | .45  |
| Forensic Risk                    | .46                                      | .53                                    | .21                                      | .14  |
| Childhood/Adolescent Misbehavior | .53                                      | .61                                    | .25                                      | .24  |

<sup>1</sup> All correlations were statistically significant at  $p < .001$  with the exception of the relationship between Narcissistic Personality Disorder and Mental Stability, which was non-significant.

**NEXT STEPS: EXPANDED SWAP/DIRE IMPLEMENTATION AND FURTHER TESTING**

As of 2010, DoE field test clinicians and eight newly trained DoE clinicians from other HRP sites are using SWAP/DIRE to evaluate: 1) employees of concern who were granted HRP eligibility despite minor red flags in their background, and are now being closely monitored by the clinical team, and 2) new referrals for cause who must receive a clinical evaluation because of a recent behavioral incident. In addition, based on the demonstrated effectiveness of SWAP/DIRE for evaluating candidate cases, DoE management is working to revise the guiding HRP policy (10 C.F.R. Part 712) to include administration of a SWAP-like instrument to all new applicants to the HRP. Although using the SWAP as a screening tool on all new candidates will result in additional clinical labor costs, DoE expects to offset additional costs (with no loss of critical information) by

reducing the frequency of administration of the currently used self-report personality disorder instruments during the continuous evaluation phase (e.g., MMPI-2-RF).

PERSEREC is also in discussions regarding potential utility and implementation of SWAP/DIRE with officials from DoD Personnel Reliability Programs. Two options under discussion include: 1) using SWAP/DIRE to evaluate PRP personnel who exhibit adverse behavior necessitating a fitness for duty evaluation, and 2) conducting a pilot test of the SWAP/DIRE at a large military base with an active mental health clinic and trained clinical staff in order to evaluate its potential utility for screening PRP candidates and current personnel.

Finally, other government agencies and departments (e.g., Department of Health and Human Services) have expressed interest in SWAP/DIRE for use with applicants to sensitive programs and/or individuals who display behaviors of concern.

## REFERENCES

- Bradley, R., Hilsenroth, M., Guarnaccia, C., & Westen, D. (2007). Relationship between clinician assessment and self-assessment of personality disorders using the SWAP-200 and PAI. *Psychological Assessment, 19*, 225-229.
- Department of Defense Independent Review. (2010). *Protecting the force: Lessons from Fort Hood*. Department of Defense. Washington, DC.
- Godes, O., & Lang, E. L. (2009). *Identifying personality disorders that are security risks: Phase I results* (TR 09-01) (FOUO). Monterey, CA: Defense Personnel Security Research Center.
- Human Reliability Program, 10 C.F.R. pt. 712 (2005).
- Lardner, R. (2011). Contractor accused of negligence in soldier deaths. *Associated Press*. Retrieved July 15, 2011, from <http://www.newsobserver.com/2011/07/12/1338582/contractor-accused-of-negligence.html>
- Marin de Avellan, L. E. (2010). *Systematizing clinician judgment of personality and risk: Validation of the SWAP-200 with forensic patients*. Unpublished Manuscript. London: University College London.
- Morey, L.C. (1991). *Personality Assessment Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.
- Shechter, O. G., & Lang, E. L. (in prep). *Identifying personality disorders that are security risks: Field test results* (FOUO). Monterey, CA: Defense Personnel Security Research Center.
- Shedler, J. (2009). *Guide to SWAP-200 interpretation: National security version* (FOUO). Aurora, CO.
- Shedler, J., & Westen, D. (1998). Refining the measurement of Axis II: A Q-sort procedure for assessing personality pathology. *Assessment, 5*, 333-353.
- Shedler, J., & Westen, D. (2004). Dimensions of personality pathology: An alternative to the Five-Factor Model. *American Journal of Psychiatry, 161*, 1743-1754.
- Tellegen, A., & Ben-Porath, Y. S. (2008). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2 Restructured Form): Technical manual*. Minneapolis, MN: University of Minnesota Press.